METHOD FOR DETECTING ILLEGAL LITERARY WORK

Publication number: JP2000215238
Publication date: 2000-08-04
Inventor: NAKADA JUNJI
Applicant: HITACHI LTD

Classification:

- international:

G06F12/14; G06F12/00; G06F21/24; G06Q10/00; G06G30/00; G06Q50/00; H04N1/387; G06F12/14; G06F12/00; G06F21/00; G06Q10/00; G06Q30/00; G06Q50/00; H04N1/387; (IPC1-7); G06F17/60:

G06F12/00; G06F12/14; H04N1/387

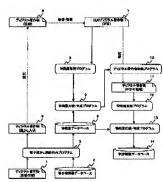
- European:

Application number: JP19990012732 19990121 Priority number(s): JP19990012732 19990121

Report a data error here

Abstract of JP2000215238

PROBLEM TO BE SOLVED: To make it possible to reficiently detect an illegal literary work on a network while maintaining a realistic cyclic period. SOL UTION: The method is provided with a means (digital work acquiring program 10) for acquiring a digital work louged as a work having high similarity, a means (feature amount acquiring program 12) for acquiring a feature amount allowed to be acquired by reading out data from the digital work and a 2nd judging means (feature amount comparing/judging program 13) for comparing the acquired feature amount with a registered feature amount with a registered feature amount and judging their similarity.



Data supplied from the esp@cenet database - Worldwide

Family list 1 family member for: JP2000215238 Derived from 1 application

Back to JP2000215

1 METHOD FOR DETECTING ILLEGAL LITERARY WORK

Inventor: NAKADA JUNJI Applicant: HITACHI LTD

EC: IPC: G06F12/14; G06F12/00; G06F21/24 (+15)

Publication info: JP2000215238 A - 2000-08-04

Data supplied from the esp@cenet database - Worldwide

(19)日本個等新庁 (JP) (12) 公開特許公報 (A)

(11)特許出願公開番号 特開2000-215238

(P2000-215238A) (43) 公開日 平成12年8月4日(2000.8.4)

(51) Int.Cl.7		裁別記号	FΙ			テーマコード(参考)
G06F	17/60		C 0 6 F	15/21	Z	5B017
	12/00	5 2 0		12/00	520E	5B049
	12/14	320		12/14	320E	5 B 0 8 2
H 0 4 N	1/387		H04N	1/387		5 C O 7 6

審査請求 未請求 請求項の数7 OL (全 5 頁)

		Di Trinisia.	Manual Manual and Control
(21)出顧番号	特願平11-12732	(71)出關人	00000j108 株式会社日立製作所
(22) 出顧日	平成11年1月21日(1999.1.21)		東京都千代田区神田駿河台四丁目 6 番地
		(72)発明者	中田 順二 神奈川県川崎市麻生区王禅寺1099番地 株 式会社日立製作所システム開発研究所内
		(74)代理人	100068504 弁理士 小川 勝男
			最終頁に続く

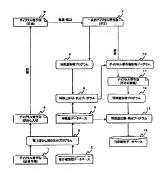
(54) 【発明の名称】 不正著作物検出方法

(57)【要約】

【課題】電子透かし技術だけに基づいた不正著作物検出 方法では、インターネット上のディジタル著作物をすべ て取得した上で、各々のディジタル著作物に透かし情報 が含まれていないかどうかを判定する必要がある。

【解決手段】類似性が高いと判定されたディジタル著作 物を取得する手段(ディジタル著作物取得プログラム1 0)と、ディジタル著作物のデータを読み取ることで初 めて取得可能となる特徴量を取得する手段(特注量取得 プログラム12)と、取得された特徴量と登録された特 徴量とを比較して類似度を判定する第二の判定手段(特 微量比較・判定プログラム13)とを設けた。

図1



【特許請求の範囲】

【請求項 1)複数のディジクル著作物の中から、あるディジクル率作物を元に作成された不正著作物を規する 不正著作物を抗力法において、ディジクル著作物的有の 複数の特徴的な値を比較することにより、不正著作物で あるかどうかを段階的に判定することを特徴とする不正 著作物検出がよ

【請求項2】ディジタル著作物固有の特徴的な値として、ファイル名称を用いることを特徴とする請求項1記 載の不正著作物検出方法。

【請求項3】ディジタル著作物固有の特徴的な値として、ファイルサイズを用いることを特徴とする請求項1 記載の不正著作物検出方法。

【請求項4】ディジタル著作物固有の特徴的な値として、ファイルのハッシュ値を用いることを特徴とする請求項1計載の不正著作物検出方法。

【請求項5】ディジクル著作物固有の特徴的な値として、ディジタル著作物のデータの一部。または全てを用いることを特徴とする請求項1記載の不正著作物検出方法。

【請求項6】ディジタル著作物固有の特徴的な値として、ディジタル著作物に埋め込まれた電子透かし情報を 用いることを特徴とする請求項1記載の不正著作物検出 方法。

【請求項7】ディジタル著作物のデータを読み取ること なく取得可能な著作物固有の特徴的な値を先に比較し、 続いてディジシル著作物のデータを読み取ることで取得 可能な著作物固和の特徴的な値を比較して不正著作物で あるか否か長限間的に判定することを特徴とする請求項 1記載の不正著作物短出方法。

【発明の詳細な説明】

[0001]

【発明の属する技術分野】本発明はディジタル著作物の 著作権保護を目的としたもので、特に不正に作成された 著作物を検出する方法に関する。

[0002]

【従来の技術】ディジタル著作物は複製や加工編集が容易なため、著作権保護技術が欠かせないものとなっている。中でも、電子透かし技術はディジクル著作物の特別にはとんど変化を与えることなく情報を書き込むことが可能なため、不正著作物の検出効果が明時されている。「0003】ところで、電子透かし技術はディジタル著作物であるか否かを電子透かしが触されたディジタル著作物であるか否かを、電子透かしが触されたディジタル著作物であるか否かをした技術だけに基づいて不正準作物を検出するためには、流通しているすべてのディジタル著作物に対して総当たりで電子透かしの有無をチェックする必要がある。「0004】 http://www.mem.co.jp/dose.htmlに記載

のデジタルコンテンツ不正利用監視センターでは、この

問題に対してインターネット検索ロボットを適用することで解決を試みている。このインターネット検索ロボットはWebページ上のディジタル著作物を機械的に次々に取得し、添かし検出システムにかけることで、電子透かしを検出している。

[0005]

【発明が解決しようとする課題】しかし、上記従来技術では、インターネット上のディジクル著作物をすべて取得した上で、各々のディジクル著作物に透かし情報が含まれていないかどうかを判定する必要がある。これは技術的には可能であるものの、インターネット上に膨大な数の地ページが存在することを考え合わせると、現実的か手段とは言えない。

【00061 現在、インターネット接続アロバイダが個人用の場合ページのために標準的に用意しているディタ容量は大体10時程度である。全世界のインターネットユーザーは少なく見積もっても数千万人は存在するから、単株に計算すると、個人用の場合ページだけで、数百形の容量になる。100kB/sの速度で他やページが取得できたとしても、数百形の場合ページを巡回するには、何百年もの時間が必要となる。これはあまりにも非現実的な数値であり、最近の場合がページの更新頻度から言っても、1~2ヶ月で一通り巡回できる必要がある。

【〇〇〇7】また、ディジタル著作物から電子透かし情報を読み出すためには、電子透かし情報を埋め込む前の 原画像が必要だったり、特殊な演算が必要だったりする。より効率的に味らページを巡回するためには、全て のディジクル著作物を取得する前や取得したディジタル 著作物の電子透かしをチェックする前にある程度の絞り 込みを行うことが求められる。

[0008]

【課題を解決するための手段】以上の課題を解決するために、本発明では、著作権を保護すべきディジタル著作物の複数の特徴量を登録・管理する手段と、インターネット上に存在するディジタル著作物のデクを添み取ることなく取得可能な特徴量とを比較して類切度を判定する第一の判定手段と、類似性が高いと判定されたディタル著作物を取得する手段と、ディジタル著作物を取得する手段と、ディジタルを指令を認め取ります。

[0009]

【発明の実験の形態】以下、図1〜図1〜図4を用いて本売明 ・ ディングル著作物販売に適用した場合の一実施例を 明する、図1は本発明の一実施例を構成するプロック 図、図2は著作権情報データベースの例、図3は特徴量 データベースの例、図4は特徴量の取得と比較・判定に 関わる部分の処理フローである。

【0010】図1において、1は著作権保護の対象とす るディジタル著作物の原著作物、2はディジタル著作物 1に関する著作権情報を管理する著作権情報データベー ス、3はディジタル著作物1に対して電子透かし処理を 施す電子透かし埋め込みプログラム、4は電子透かし処 理の施されたディジタル著作物、5はディジタル著作物 4の特徴量を管理する特徴量データベース、6はディジ タル著作物4を販売した結果として購入者側に作成され る正規のディジタル著作物、7はディジタル著作物6を 複製することにより作成され、ネットワーク上に不正に 掲載される二次的ディジタル著作物、8はネットワーク 上のディジタル著作物の特徴量を取得する特徴量取得プ ログラム、9は特徴量取得プログラム8が出力する特徴 量と特徴量データベース5の内容とを比較し、類似した ものがないかどうかを判定する特徴量比較・判定プログ ラム、10は特徴量比較・判定プログラム9の判定結果 に従って呼び出されて、二次的ディジタル著作物7を随 時取得するディジタル著作物取得プログラム、11はデ ィジタル著作物取得プログラム10によって取得された 不正の疑義のあるディジタル著作物、12は取得したデ ィジタル著作物11の特徴量を取得する特徴量取得プロ グラム 13 は特徴量取得プログラム 12 が出力する特 徴量と特徴量データベース5の内容とを比較し、該当す るものがないかどうかを判定する特徴量比較・判定プロ グラム 14は特徴量比較・判定プログラム13の判定 結果を格納する判定結果データベースである。

【0011】まず、電子透かし埋め込みプログラム3を 用いて、ディジタル著作物1に対して著作権情報データ ベース2から読み出した情報を透かし情報として埋め込 む、著作権情報データベース2を構成するテーブルとし ては図2に示すような3つのテーブル、すなわち著作物 テーブル15、購入者テーブル16、売上テーブル17 が考えられるが、透かし情報として埋め込むものは著作 物テーブル15の各要素を一意に識別する著作物IDだけ でもよいし、売上テーブル17の各要素を一意に識別す る売上IDだけでもよい。または、著作物テーブル15に 含まれる著作物名称や購入者テーブル16に含まれる購 入者名称などを直接透かし情報として利用してもよい。 【0012】続いて、電子透かし埋め込みプログラム3 は、透かし情報を埋め込んだディジタル著作物4を作成 するとともに、ディジタル著作物4に関する特徴量を特 徴量データベース5に記録する。特徴量データベース5 を構成するテーブルとしては図3に示すような要素をも つ特徴量テーブル18が考えられる。

【0013】特徴量テーブル18は、特徴量を一窓に識別する特徴量ID、特徴量の対象となるディジクル著作物を一窓に識別する著作物ID、ディジクル著作物のファイル名称、ファイルサイズ、ファイルのハッシュ値、ハッシュ値を得るのに利用したハッシュ関数、ファイルの実体、透かし情報、どのような情報を透かし情報として用

いたかを表すための透かし情報の種類、から構成される。

【0014】ディジタル者中郷々は正規の販売手続きを経て開入者へ渡され、購入者のもとでディジタル者作物 6として利用される。私が文利用の範囲にとどまる複製 は特に法律で規制するものではないが、ネットワーク上へ掲載された場合は不正な二次的ディジタル者作物で 12をのようなのようなのようなのようなのようなのは、これを峻別する必要がある。この峻別作業から不正に複製された著作物と判定するまでの処理フローについて図4を用いて認明する。

(S001) 特徴量取得プログラム8は、ネットワーク上の あらゆるサイトを自動的に探索するロボットプログラム で、特徴量としてファイルの名称やファイルのサイズを ファイルのURLとともに取得する。

【0015】(5002)特徴重比較・判定プログラム9 は、特徴重視等プログラムおが取得したファイル名称や ファイルが大と合数するしコードが特徴量データベー ス5にないかどうかを検索する。ここで、ファイル名称 はファイルの内容を変えずに簡単に変更できるため、フ ァイル名称の中でファイルの種別を表す拡張子部分のみ を比較してもよい。

【0016】(S003)特徴量データベース5に含まれるファイルサイズとファイルを称の組と同一または類似の特徴量を持つホットワーク上のディジタル著作物を、先に取得したIRLからディジタル著作物取得プログラム10を用いて取得する。

【0017】(S004)取得したディジタル著作物11の ハッシュ値を、特徴量データベース5の該当するレコー ドに登録されているハッシュ関数を用いて算出する。

【0018】(\$005) 算出したハッシュ値が、特徴量デ ータベース5の該当するレコードと一致するかどうかを 比較判定する。一致する場合は(\$006) に進む。一致し ない場合は処理を終了する。

【0019】(\$006)ディジタル著作物11と特徴量デ ータベースの該当するレコードに合まれるファイル実体 とを比較する。一致する場合は(\$007)に進む。一致し ない場合は処理を終了する。

【0020】(S007) ディジタル著作物11に含まれる 透かし情報と特徴量データベースの該当するレコードに 含まれる透かし情報とを比較判定する。

【0021】(S008) 判定結果を判定結果データベース 14に記録して処理を終了する。判定結果には、透かし 情報を除いた特徴量が特徴量データベース5に登録され たものと一致するネットワーク上のディジタル著作物の 所在に関する情報が格納されることになる。

[0022]

【発明の効果】本発明によると、電子透かしを用いた不正著作物検出方法において、ネットワーク上の全てのデ

ィジタル著作物を取得する前にある程度の絞り込みを行 うことが可能となるため、現実的な巡回期間を維持した ままで、効果的にネットワーク上の不正著作物の検出が 可能となる。

【0023】例えば、平均的なディジタル著作物のファ イルサイズを100kBとした場合、そのファイル名称やフ ィルサイズなどの特徴量の情報は高々0.1kB程度で記 速可能なため、ネットワーク上の著作物を駆ける速度 が全ての著作物を取得する場合に比べて1000倍になる。 つまり、全ての著作物を取得したのでは何百年もかかる と考えられる巡回期間が、数ヶ月程度まで短縮できると 言える。

【0024】もちろん、数ヶ月もの間隔を空けたのでは、先にも極いたように最近の地かページの更新頻度から言ってもまだ不十分であるが、この程度の性能不足ならば数台のサーバ機を設置して並列処理することで、十分に現実的なニーズに即した巡回が可能である。

【図面の簡単な説明】

選手掛か! 乗め込みプログラム

要件報情報データペール

【図1】本発明の一実施例を構成するプロック図であ 【図1】 **3.**

【図2】著作権情報データベースの例を示す図である。 【図3】特徴量データベースの例を示す図である。

【図4】特徴量の取得と比較・判定に関わる部分の処理 フローである。

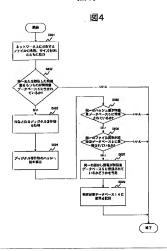
【符号の説明】

[図2]

図1 図2
| FRIND | BYRCE | BEE ARE | FRIT / 7/2) A FREE | FRIT / 7/2 A FRIT

利度対象データベース





フロントページの続き

ドターム(参考) 58017 AA06 BA05 BA07 B803 CA16 58049 AA05 B826 CC22 CC23 DD05 B803 B805 EE56 FF09 GG04 GG07 GG09 58082 AA13 BA07 EA09 GA02 GC03 GC04 50076 AA14 BA06